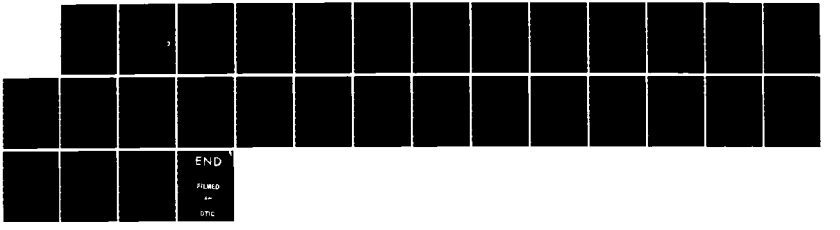
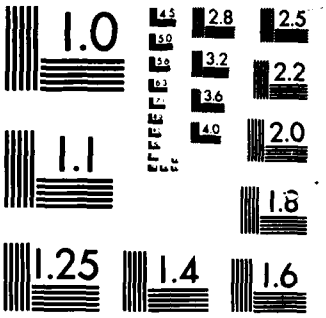
AD-A163 329 GENERALIZED ORDER STATISTICS BAHADUR REPRESENTATIONS AND SEQUENTIAL NONPA..(U) JOHNS HOPKINS UNIV BALTIMORE MD DEPT OF MATHEMATICAL SCIENCES.. J CHOUDHURY ET AL. 1/1

UNCLASSIFIED OCT 85 TR-445 N00014-79-C-0801 F/G 12/1 NL






MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

DEPARTMENT OF MATHEMATICAL SCIENCES
The Johns Hopkins University
Baltimore, Maryland 21218

# GENERALIZED ORDER STATISTICS, BAHADUR REPRESENTATIONS, AND SEQUENTIAL NONPARAMETRIC FIXED-WIDTH CONFIDENCE INTERVALS



by

J. Choudhury[1] and R.J. Serfling



Technical Report No. 445
ONR Technical Report No. 85-4
October, 1985



Research supported by the U.S. Department of Navy under Office of Naval Research Contract No. N00014-79-C-0801. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[1]University of Baltimore, Baltimore, Maryland 21218

(-1-)

## ABSTRACT


GENERALIZED ORDER STATISTICS, BAHADUR REPRESENTATIONS, AND SEQUENTIAL NONPARAMETRIC FIXED-WIDTH CONFIDENCE INTERVALS

Let $X_1,\ldots,X_n$ be an i.i.d. sample from df F, let $H_F$ be the df of $h(X_1,\ldots,X_m)$, based on a given "kernel" $h(x_1,\ldots,x_m)$, and consider confidence interval estimation of a parameter of the form $H_F^{-1}(p)$. This paper introduces confidence intervals formed by a pair of "generalized order statistics," develops Bahadur-type representation theory for these order statistics, and constructs corresponding sequential fixed-width confidence interval procedures. Previous work of Bahadur (1966) and Geertsema (1970) is sharpened and extended.


Keywords: nonparametric statistics; asymptotic theory.

1. Introduction. In this paper we introduce a notion of generalized order statistics, develop relevant asymptotic theory, and apply the results to characterize the convergence properties of a class of sequential nonparametric fixed-width confidence interval procedures. Previous work of Bahadur (1966) and Geertsema (1970) is broadly extended, in close connection with ideas introduced in Serfling (1984).

Let $X_1,\ldots,X_n$ be independent random variables having common distribution function (df) F. (More generally, the $X_i$'s may be random elements of an arbitrary space.) Let h be a function from $\mathbb{R}^m$ to $\mathbb{R}$ and denote by $H_F$ the df of $h(X_1,\ldots,X_m)$. Estimation of parameters of F which are expressible as $T(H_F)$, where $T(\cdot)$ is a general form of L-functional, has been considered by Serfling (1984) and Janssen, Serfling, and Veraverbeke (1984). Here we confine attention to the special case of *quantile* L-functionals and hence to parameters of the form $H_F^{-1}(p), 0 < p < 1$, and we investigate nonparametric confidence intervals formed by a pair of the "generalized order statistics"

$$W_{n,1} \le \cdots \le W_{n,n_{(m)}}, \tag{1.1}$$

the ordered values of $h(X_{i_1},\ldots,X_{i_m})$ taken over the $n_{(m)} = n(n-1)\ldots(n-m+1)$ m-tuples $(i_1\ldots,i_m)$ of distinct elements from $\{1,\ldots,n\}$. For any such parameter $H_F^{-1}(p)$, the relevant sequential confidence interval will be given by

---

*AMS* 1980 *subject classifications:* Primary 60F15, Secondary 62L10

*Key words and phrases:* order statistics, Bahadur representation, sequential, nonparametric, fixed-width confidence intervals

(1.2) $$(W_{N,a(N)},\ W_{N,b(N)}),$$

where the "rank functions" a(n) and b(n) are selected so that (1.2) has specified asymptotic coverage probability for N = n (nonrandom) $\to \infty$, and where N is a random sample size selected for (1.2) to have specified fixed width.

For the case h(x) = x, (1.1) gives the usual order statistics of the sample and (1.2) represents a sequential version of the classical method of giving a nonparametric distribution-free confidence interval for a quantile $F^{-1}(p)$. For $p = \frac{1}{2}$, this sequential approach and also the one based on $h(x_1,x_2) = (x_1 + x_2)/2$ were introduced and investigated by Geertsema (1970) as competing approaches, in the case of symmetric df F, for estimation of the location parameter $H_F^{-1}(\frac{1}{2}) = F^{-1}(\frac{1}{2})$. In particular, with N = N(d) designed for (1.2) to have width 2d, and with a(n),b(n) designed to yield asymptotic coverage probability $1-2\alpha$, Geertsema characterized the convergence rate of N(d) to $\infty$ and established the convergence of the coverage probability of (1.2) to $1 - 2\alpha$, as $d \to 0$.

In the present paper we consider the behavior of the sequential confidence interval (1.2) and the random sample size N(d) for the general case of an arbitrary "kernel" $h(x_1,\ldots,x_m)$. This generality entails the complication, fortuitously absent in the two special cases treated by Geertsema, that the functions a(n),b(n) used in (1.2) may be random. Consequently, it becomes necessary to extend the representation theorem of Bahadur (1966) for central order statistics $X_{n,k_n}$, where $k_n/n \to p$, $0 < p < 1$, not only to the case of our generalized order statistics $W_{n,k_n}$, where $k_n/n_{(m)} \to p$, $0 < p < 1$, but also to the

case that $k_n$ is random.

In Section 2 we provide some convergence results on the empirical df $H_n$ and quantile function $H_n^{-1}$ associated with the $W_{n,k}$'s, and we use these results to define appropriate rank functions a(n),b(n) for use in (1.2). This empirical df is defined by

$$(1.3) \qquad H_n(y) = \frac{1}{n_{(m)}} \sum \mathbb{1}\{h(X_{i_1},\ldots,X_{i_m}) \le y\}, \quad -\infty < y < \infty,$$

where the sum is taken over the $n_{(m)}$ m-tuples $(i_1,\ldots,i_m)$ of distinct elements from $\{1,\ldots,n\}$. Clearly, $H_n(y)$ is an unbiased estimator of $H_F(y)$, and $H_n^{-1}(p)$ provides an estimator of $H_F^{-1}(p)$ analogous to the usual sample quantile as estimator of $F^{-1}(p)$.

Our extended Bahadur representation for $W_{n,k_n}$ and related results are developed in Section 3. As special cases, we obtain the results of Bahadur (1966) for the case $h(x) = x$ and $k_n$ nonrandom and of Geertsema (1970) for the case $h(x_1,x_2) = (x_1 + x_2)/2$ and $k_n$ nonrandom, under relaxations of their regularity conditions on F. The results of Section 3 and in part Section 2 are of general interest, besides their applications in this paper.

Section 4 carries out general application to the class of sequential nonparametric confidence intervals of form (1.2). The two examples treated by Geertsema (1970) are obtained as special cases, but under weaker regularity conditions on F.

The random a(n),b(n) used in defining (1.2) are constructed in terms of an estimator for a parameter appearing in the asymptotic distribution of the random variable $H_n(H_F^{-1}(p))$. It is necessary for

our theory in Section 4 that the estimator be strongly consistent. Such an estimator is developed in Section 5.

We conclude this introduction with selected examples to which the methods of Section 4 may be applied.

(i) *location estimation:* One may view the cases considered by Geertsema (1970) as two special cases of the class of kernels given by $h(x_1,\ldots,x_m) = (x_1 + \ldots + x_m)/m$, for $m = 1,2,3,\ldots$. For symmetric F, the corresponding parameters $H_F^{-1}(\frac{1}{2})$ all reduce to $F^{-1}(\frac{1}{2})$, so that the corresponding estimators $T_{nm}$ given by $H_n^{-1}(\frac{1}{2})$ are competitors for the same goal. A comparative study of these estimators for $m = 1,\ldots,5$ has been carried out in Choudhury (1984), on the basis of which a particular choice of m may be selected and the results of Section 4 utilized to provide associated sequential fixed-width confidence intervals for $F^{-1}(\frac{1}{2})$.

More generally, let us consider the kernel $h(x_1,\ldots,x_m) = \sum_{i=1}^{m} \alpha_i x_i$, with $\sum_{i=1}^{m} \alpha_i = 1$ (but the $\alpha_i$'s otherwise unrestricted). For symmetric F, the corresponding parameter $H_F^{-1}(\frac{1}{2})$ reduces in each case to $F^{-1}(\frac{1}{2})$, but the corresponding estimators $H_n^{-1}(\frac{1}{2})$ differ and therefore are competitors. The case $m = 2$ was introduced by Maritz, Wu and Staudte (1977) and studied as a special case of the class of $M_2$-estimators of Huber (1964), by switching to the closely related estimators $H_{F_n}^{-1}(\frac{1}{2})$, where $F_n$ is the usual sample df. They established asymptotic normality and examined asymptotic relative efficiencies, among other aspects. However, by noting that the statistics $H_n^{-1}(\frac{1}{2})$ are special cases of the generalized L-statistics of Serfling (1984), we can treat them

directly and obtain not only the relevant asymptotic convergence theory but also (by our Section 4) associated sequential fixed-width confidence intervals. From the numerical studies of Maritz, Wu and Staudte (1977) for the case $m = 2$ and $\alpha_1$ arbitrary, and of Choudhury (1984) for the case $\alpha_1 = \ldots = \alpha_m = 1/m$, with $m$ arbitrary, it is found that the classical median and Hodges-Lehmann estimators can be successfully competed with in various situations by the estimators $H_n^{-1}(\frac{1}{2})$ corresponding to choices of $\alpha_i$ even outside the interval $[0,1]$ and choices of $m > 2$. It would be of interest to extend these two previous studies to the case of arbitrary $m, \alpha_1, \ldots, \alpha_m$ subject to $\sum_{i=1}^{m} \alpha_i = 1$.

(ii) *spread estimation:* Included among various measures of spread discussed by Bickel and Lehmann (1979) is the median of the distribution of $|X_1 - X_2|$, where $X_1, X_2$ are independent r.v.'s having df $F$. In our context, this is a "generalized L-functional" parameter $H_F^{-1}(\frac{1}{2})$, where $H_F$ is based on the kernel $h(x_1,x_2) = |x_1 - x_2|$. This can be estimated by the generalized L-statistic $H_n^{-1}(\frac{1}{2})$ and sequential fixed-width confidence intervals can be developed by our Section 4 results. (It would be of interest to consider a general class of spread measures of this type, defined by $H_F^{-1}(\frac{1}{2})$ with $H_F$ based on a kernel of form

$$h(x_1,\ldots,x_m) = \left|\sum_{i=1}^{m} \beta_i x_i\right|, \text{ where } \sum_{i=1}^{m} \beta_i = 0.)$$

(iii) *regression slope estimation:* Consider the simple linear regression model $Y_i = \alpha + \beta X_i + \varepsilon_i$, with $\{\varepsilon_i\}$ i.i.d. r.v.'s independent

of $\{X_i\}$, and $\{X_i\}$ a sequence of *random* regressors. Let $\underset{\sim}{F}$ denote the common df of the mutually independent pairs $(X_i,Y_i)$, $1 \le i \le n$, and let $H_{\underset{\sim}{F}}$ denote the cdf of $h((X_1,Y_1),(X_2,Y_2))$, where $h((x_1,y_1),(x_2,y_2)) = (y_2 - y_1)/(x_2 - x_1)$. Then clearly a natural estimator of the parameter $\beta$ is the median of the ratios $(Y_i - Y_j)/(X_i - X_j)$, i.e., the estimator $\hat{\beta} = H_n^{-1}(\frac{1}{2})$. This is a version (for random regressors) of the well-known estimator of Theil (1950). Using the results of our Section 4, we can provide sequential fixed-width confidence intervals associated with this estimator.

2. Convergence results for $H_n$ and $H_n^{-1}$ and other preliminaries. We note that, for each fixed y, $H_n(y)$ is a U-statistic based on the kernel

$$g_y(x_1,\ldots,x_m) = \mathbb{1}\{h(x_1,\ldots,x_m) \le y\}, \quad (x_1,\ldots,x_m) \in \mathbb{R}^m.$$

Consequently, by standard results on U-statistics (e.g., Serfling (1980), Chapter 5), we have *strong convergence* and *asymptotic normality*:

$$H_n(y) \xrightarrow{a.s.} H_F(y), \quad n \to \infty, \tag{2.1}$$

and

$$n^{\frac{1}{2}}[H_n(y) - H_F(y)] \xrightarrow{d} N(0, m^2\sigma_y^2), \tag{2.2}$$

where $\sigma_y^2 = \mathrm{Var}_F\{g_{y1}(X)\}$, with $g_{y1}(x) = \mathrm{Var}\{E\{\sum_A g_{y1}(X_{i_1},\ldots,X_{i_m})/m! \mid X_1 = x\}$, where $\sum_A$ denotes summation over all permutations of $(1,\ldots,m)$.

We shall be applying (2.2) with $y = H_F^{-1}(p)$, in which case a key

parameter of concern will be

$$\zeta_p = \sigma^2_{H_F^{-1}(p)}. \tag{2.3}$$

Other convenient notation will be $\xi_p = H_F^{-1}(p)$ and $\hat{\xi}_{pn} = H_n^{-1}(p)$.

By (2.1) and an argument similar to the proof of strong convergence of the classical sample quantile (e.g., Serfling (1980), §2.3), we obtain *strong convergence* of $\hat{\xi}_{np}$,

$$\hat{\xi}_{pn} \xrightarrow{a.s.} \xi_p, \tag{2.4}$$

under the condition that $\xi_p$ is the unique solution of $H_F(y-) \le \xi_p \le H_F(y)$. Also, by Serfling (1984), we have *asymptotic normality* of $\hat{\xi}_{pn}$,

$$n^{\frac{1}{2}}(\hat{\xi}_{pn} - \xi_p) \xrightarrow{d} N(0, m^2\zeta_p/h_F^2(\xi_p)), \tag{2.5}$$

where it is assumed that $H_F$ has density $h_f$ positive at $\xi_p$.

One could use (2.5) as a basis for construction of confidence intervals for $\xi_p$, but this would entail estimation of both $\zeta_p$ and $h_F(\xi_p)$. Our approach based on the generalized order statistics (1.1) eliminates estimation of the latter parameter.

Let us now formulate the rank functions a(n),b(n) used in defining the interval (1.2). First, we note that for integer $k_n$ we have

$$P\{W_{n,k_n} \le \xi_p\} = P\{H_n(\xi_p) \ge k_n/n_{(m)}\}$$

$$= P\{n^{\frac{1}{2}}(H_n(\xi_p) - H_F(\xi_p)) \geq n^{\frac{1}{2}}(\frac{k_n}{n_{(m)}} - H_F(\xi_p))\}. \tag{2.6}$$

If $k_n$ is defined by (with $\Phi$ denoting the standard normal cdf)

$$\frac{k_n}{n_{(m)}} = p + \frac{\Phi^{-1}(1-\alpha)m\zeta_p^{\frac{1}{2}}}{n^{\frac{1}{2}}} + o(n^{-\frac{1}{2}}),\ n \to \infty, \tag{2.7}$$

then by (2.2) and (2.6) it follows that

$$P\{W_{n,k_n} \leq \xi_p\} \to \alpha,\ n \to \infty. \tag{2.8}$$

Moreover, (2.8) remains true if $\zeta_p$ is replaced by a consistent estimator $\hat{\zeta}_{pn}$ in (2.7). By similar arguments, if $\zeta_p^{\frac{1}{2}}$ is replaced by $-\zeta_p^{\frac{1}{2}}$ or $-\hat{\zeta}_{pn}^{\frac{1}{2}}$ in (2.7), then

$$P\{W_{n,k_n} \geq \xi_p\} \to \alpha,\ n \to \infty. \tag{2.9}$$

On this basis we define integers $a(n)$, $b(n)$ by

$$\frac{a(n)}{n_{(m)}} = p - \frac{\Phi^{-1}(1-\alpha)m\zeta_p^{\frac{1}{2}}}{n^{\frac{1}{2}}} + o_p(n^{-\frac{1}{2}}) \tag{2.10}$$

and

$$\frac{b(n)}{n_{(m)}} = p + \frac{\Phi^{-1}(1-\alpha)m\zeta_p^{\frac{1}{2}}}{n^{\frac{1}{2}}} + o_p(n^{-\frac{1}{2}}) \tag{2.11}$$

and assert that

$$P\{(W_{n,a(n)}, W_{n,b(n)}) \text{ contains } \xi_p\} \to 1 - 2\alpha,\ n \to \infty. \tag{2.12}$$

In practice $\zeta_p$ is unknown and must be estimated (consistently) to obtain a(n),b(n) satisfying (2.10), (2.11). Moreover, for the sequential analogue of (2.12) obtained in Section 4, we shall need $\hat{\zeta}_{pn}$ to be strongly consistent. A suitable such estimator is developed in Section 5.

For certain special cases of kernel h, the "parameter" $\zeta_p$ does not depend on F. For example, in the case $h(x) = x$ we have $\zeta_p = p(1-p)$; in the case $h(x_1,x_2) = (x_1 + x_2)/2$ and $p = \frac{1}{2}$, we have $\zeta_{\frac{1}{2}} = 1/12$. (Thus Geertsema (1970) did not have to deal with random versions of the functions a(n),b(n).) In general, however, $\zeta_p$ depends upon F. For example, in the case $h(x_1,\ldots,x_m) = (x_1+\ldots+x_n)/m$, and $F(x) = F_0(x - \xi_{\frac{1}{2}})$, and $m \geq 2$, we have

$$\zeta_{\frac{1}{2}}(F) = \mathrm{Var}_F\{F_0^{(m-1)}(X)\},$$

where $F_0^{(k)}$ denotes the k-th order convolution of $F_0$. For $m \geq 3$, this parameter may be seen to depend upon F.

3. Generalized order statistics and Bahadur representation theory. We consider the order statistics $W_{n,k}$ defined by (1.1). Our first result provides an a.s. error bound for $W_{n,k_n}$ as an estimator of $\xi_p = H_F^{-1}(p)$, when $k_n/n_{(m)}$ converges to p at a suitably fast a.s. rate. (This generalizes and sharpens Lemma 2.5.4C of Serfling (1980), given for the classical order statistics.)

LEMMA 3.1. *Let* $0 < p < 1$. *Suppose that* $H_F$ *is differentiable at* $\xi_p$ *with* $H_F'(\xi_p) = h_F(\xi_p) > 0$. *Let* $\{k_n\}$ *be a sequence of positive*

*integer-valued r.v.'s* $(1 \le k_n \le n_{(m)})$ *such that*

$$k_n/n_{(m)} - p = o(\varepsilon_n), n \to \infty, \text{ a.s.,} \tag{3.1}$$

*where* $\{\varepsilon_n\}$ *is a sequence of constants tending to 0 with*

$$\varepsilon_n^2 n(\log n)^{-1} h_f^2(\xi_p)/m > c_0 > 1, \text{ all } n \text{ sufficiently large.} \tag{3.2}$$

*Then a.s.*

$$|W_{n,k_n} - \xi_p| \le \varepsilon_n, \text{ all } n \text{ sufficiently large.} \tag{3.3}$$

PROOF. We must show that a.s.

$$W_{n,k_n} \le \xi_p + \varepsilon_n, \text{ all n sufficiently large,} \tag{3.4}$$

and

$$W_{n,k_n} \ge \xi_p - \varepsilon_n, \text{ all n sufficiently large.} \tag{3.5}$$

We shall prove (3.4), the proof of (3.5) being similar. Note that (3.4) is equivalent to

$$H_n(\xi_p + \varepsilon_n) - H_F(\xi_p + \varepsilon_n) \ge \frac{k_n}{n_{(m)}} - H_F(\xi_p + \varepsilon_n), \text{ all large n.} \tag{3.6}$$

Now, by application of a probability inequality of Hoeffding (1963) (or see Serfling (1980), p. 201), we have

$$P\{H_n(\xi_p + \varepsilon_n) - H_F(\xi_p + \varepsilon_n) < t\} \le e^{-2[n/m]t^2}, t < 0, n \ge m. \tag{3.7}$$

For $t = -2^{-\frac{1}{2}} h_F(\xi_p)\varepsilon_n$, the LHS of (3.7) is seen to be $0(n^{-c_0})$, whence by the Borel-Cantelli Lemma we have that a.s.

$$(3.8) \qquad H_n(\xi_p + \varepsilon_n) - H_F(\xi_p + \varepsilon_n) > -2^{\frac{1}{2}} h_F(\xi_p)\varepsilon_n, \text{ all large } n.$$

On the other hand, by (3.1) and Young's form of Taylor's Theorem (e.g., see Serfling (1980), p. 45), we have a.s.

$$(3.9) \qquad \frac{k_n}{n_{(m)}} - H_F(\xi_p + \varepsilon_n) = H_F(\xi_p) - H_F(\xi_p + \varepsilon_n) + o(\varepsilon_n)$$

$$= -h_F(\xi_p)\varepsilon_n + o(\varepsilon_n)$$

$$< -2^{\frac{1}{2}} h_F(\xi_n)\varepsilon_n, \text{ all large } n.$$

Thus (3.6) holds a.s. □

Next we provide a modulus-of-continuity-type result for the empirical process $H_n(\cdot) - H_F(\cdot)$. This strengthens an earlier version given by Serfling and Thornton (1981) and also generalizes Lemma 2.5.4E of Serfling (1980).

LEMMA 3.2. *Let* $0 < p < 1$ *and put* $\xi_p = H_F^{-1}(p)$. *Suppose that* $H_F'$ *is bounded in a neighborhood of* $\xi_p$, *with* $H_F'(\xi_p) = h_F(\xi_p) > 0$. *Let* $\{a_n\}$ *be a sequence of constants tending to* 0 *with*

$$(3.10) \qquad a_n n^{\frac{1}{2}} (\log n)^{-\frac{1}{2}} > \Delta > 0.$$

*Put*

$$(3.11) \qquad T_{pn} = \sup_{|y| \le a_n} \left| [H_n(\xi_p + y) - H_n(\xi_p)] - [H_F(\xi_p + y) - H_F(\xi_p)] \right|.$$

*Then a.s.*

$$T_{pn} = O(a_n^{\frac{1}{2}} n^{-\frac{1}{2}} (\log n)^{\frac{1}{2}}), \quad n \to \infty \tag{3.12}$$

PROOF. The argument used for Lemma 2.5.4E of Serfling (1980) carries through, with the use of Theorem 5.6.1A of Serfling (1980) in place of Lemma 2.5.4A of Serfling (1980), and the use of Lemma 3.1 (above) in place of Lemma 2.5.4C of Serfling (1980). □

The next result provides a Bahadur-type representation for $W_{n,k_n}$.

THEOREM 3.1. *Let* $0 < p < 1$ *and put* $\xi_p = H_F^{-1}(\xi_p)$. *Suppose that* $H_F$ *is twice differentiable at* $\xi_p$, *with* $H_F'(\xi_p) = h_f(\xi_p) > 0$. *Let* $\{k_n\}$ *be a sequence of positive integer-valued r.v.'s* $(1 \le k_n \le n_{(m)})$ *satisfying* (3.1) *and* (3.2). *Then a.s.*

$$W_{n,k_n} = \xi_p + \frac{k_n/n_{(m)} - H_n(\xi_p)}{h_F(\xi_p)} + O(\max\{\varepsilon_n^2, \varepsilon_n^{\frac{1}{2}} n^{-\frac{1}{2}} (\log n)^{\frac{1}{2}}\}), n \to \infty. \tag{3.13}$$

PROOF. Under the assumptions of Theorem 3.1, Lemma 3.1 is applicable, from which we have a.s.

$$\left| W_{n,k_n} - \xi_p \right| \le \varepsilon_n, \quad n \to \infty. \tag{3.14}$$

Since $\varepsilon_n > \varepsilon_n^2$ for all n sufficiently large, Lemma 3.2 (with $\varepsilon_n$ in place of $a_n$) is applicable, whence using (3.14) we have

$$[H_n(W_{n,k_n}) - H_n(\xi_p)] - [H_F(W_{n,k_n}) - H_F(\xi_p)] = O(\varepsilon_n^{\frac{1}{2}} n^{-\frac{1}{2}} (\log n)^{\frac{1}{2}}), n \to \infty. \tag{3.15}$$

Now, by Young's form of Taylor's Theorem (Serfling (1980), p. 45) and (3.14) we write, a.s.,

$$(3.16)\quad H_F(W_{n,k_n}) = H_F(\xi_p) + (W_{n,k_n} - \xi_p)h_F(\xi_p) + (W_{n,k_n} - \xi_p)^2 h_F'(\xi_p)/2! + o(\varepsilon_n^2),\ n \to \infty .$$

Since $H_n(W_{n,k_n}) = k_n/n_{(m)}$ a.s., (3.14), (3.15) and (3.16) yield (3.13). □

This result extends the classical result of Bahadur (1966) (or see Serfling (1980), p. 91) to the case of random $k_n$ and generalizes to the $W_{n,k_n}$ as well. Also, the regularity condition of Bahadur (1966) that F" (in our general context, $H_F$") be bounded in a neighborhood of $\xi_p$ is slightly relaxed. Indeed, one can further relax this regularity condition and obtain the following useful variant of Theorem 3.1.

LEMMA 3.3. *Let* $0 < p < 1$ *and put* $\xi_p = H_F^{-1}(\xi_p)$. *Suppose that* $H_F'$ *is bounded in a neighborhood of* $\xi_p$, *with* $H_F'(\xi_p) = h_F(\xi_p) > 0$. *Let* $\{k_n\}$ *be as in Theorem* 3.1. *Then a.s.*

$$(3.17)\quad W_{n,k_n} = \xi_p + \frac{k_n/n_{(m)} - H_n(\xi_p)}{h_F(\omega_n^*)} + O(\varepsilon_n^{1/2} n^{-1/2} (\log n)^{1/2}),\ n \to \infty ,$$

*where* $\omega_n^*$ *lies between* $\xi_p$ *and* $W_{n,k_n}$.

(The proof is similar to that of Theorem 3.1.)

LEMMA 3.3 will be used to advantage in Section 4. One can obtain a further variant of Theorem 3.1, involving further relaxation of the regularity condition on $H_F$ but yielding a slower rate in (3.17) and only in probability instead of almost surely. This is analogous to a variant of Bahadur's result given by Ghosh (1971). However, for the results of Section 4, the version given by Lemma 3.3 is needed.

4. Sequential nonparametric fixed-width confidence intervals. Let integers a(n) and b(n) be defined via the formulas of (2.10) and (2.11) with $\zeta_p$ replaced by an estimator $\hat{\zeta}_{pn}$. Let $N(=N(d))$ be the smallest integer $n \geq n_0$ for which

$$W_{n,b(n)} - W_{n,a(n)} \leq 2d, \tag{4.1}$$

for some specified $d > 0$ and $n_0 \geq 1$. We consider the sequential 2d-width confidence interval

$$(W_{N,a(N)}, W_{N,b(N)}) \tag{4.2}$$

for estimation of $\xi_p = H_F^{-1}(p)$. The key properties of this sequential confidence interval procedure are given by the following result.

THEOREM 4.1. *Suppose that* $H_F' = h_F$ *is positive and Lipschitz of order* $\Delta$ *at* $\xi_p$, *for some* $\Delta > 0$. *Let* $\hat{\zeta}_{pn}$ *be a strongly consistent estimator of* $\zeta_p$. *Then the sequential fixed-width confidence interval procedure defined by* (4.2), *and the random sample size* N *required by*

*this procedure, have the properties*

(a) N *is well-defined for all* $d > 0$, $N(=N(d))$ *is a nondecreasing function of* d *as* d *decreases to* 0, $\lim_{d\to 0} N(d) = \infty$ a.s., *and* $\lim_{d\to 0} E(N) = \infty$.

(b) $\lim_{d\to 0} Nd^2 = [\Phi^{-1}(1-\alpha)]^2 m^2 \zeta_p / h_F^2(\xi_p)$ *a.s.*.

(c) $\lim_{d\to 0} P\{W_{N,a(N)} \leq \xi_p \leq W_{N,b(N)}\} = 1 - 2\alpha$.

*Under the additional assumption*

$$\sup_{n \geq n_0} E(\hat{\zeta}_{pn}^{\beta}) < \infty \text{ for some } \beta > 1, \tag{4.3}$$

*we have also*

(d) $\lim_{d\to 0} E(N)d^2 = [\Phi^{-1}(1-\alpha)]^2 m^2 \zeta_p / h_F^2(\xi_p)$; *and* $E(N) < \infty, d > 0$.

We first establish three lemmas needed for the proof of this theorem.

LEMMA 4.1. *Let integers* $a(n), b(n)$ *be defined via the formulas of* (2.10), (2.11) *with* $\zeta_p$ *replaced by a strongly consistent estimator* $\hat{\zeta}_{pn}$. *Let* $H_F$ *satisfy the assumptions of Theorem* 4.1. *Then a.s.*

$$n^{\frac{1}{2}}(W_{n,b(n)} - W_{n,a(n)}) \to 2\Phi^{-1}(1-\alpha) m \zeta_p^{\frac{1}{2}} / h_F(\xi_p), \quad n \to \infty \tag{4.4}$$

PROOF. By the strong consistency of $\hat{\zeta}_{pn}$, it follows that a.s.

$$a(n)/n_{(m)} = p + o(\varepsilon_n), \quad n \to \infty, \tag{4.5}$$

and

(4.6) $$b(n)/n_{(m)} = p + o(\varepsilon_n),\ n \to \infty,$$

where, with $c_0 > 1$,

(4.7) $$\varepsilon_n = (c_0 \log n)^{\frac{1}{2}} m^{\frac{1}{2}} n^{-\frac{1}{2}} / h_F(\xi_p).$$

Hence Lemma 3.3 applies twice, with $k_n$ given by $a(n)$ and $b(n)$, and yields a.s.

(4.8) $$n^{\frac{1}{2}}(W_{n,b(n)} - W_{n,a(n)}) = n^{\frac{1}{2}}[p - H_n(\xi_p)](h_F^{-1}(\omega_{n2}^*) - h_F^{-1}(\omega_{n1}^*))$$
$$+ \Phi^{-1}(1-\alpha) m \hat{\zeta}_{pn}^{\frac{1}{2}} (h_F^{-1}(\omega_{n2}^*) + h_F^{-1}(\omega_{n1}^*))$$
$$+ O(\varepsilon_n^{\frac{1}{2}} (\log n)^{\frac{1}{2}}),\ n \to \infty,$$

where $\omega_{n1}^*$ lies betwen $\xi_p$ and $W_{n,a(n)}$, and $\omega_{n2}^*$ between $\xi_p$ and $W_{n,b(n)}$. By Lemma 3.1, and the Lipschitz assumption on $h_F$, we have a.s.

$$(h_F^{-1}(\omega_{n2}^*) - h_F^{-1}(\omega_{n1}^*)) = O(\varepsilon_n^{\Delta}) = O(n^{-\Delta/2} (\log n)^{\Delta/2}),\ n \to \infty.$$

Now, by the law of iterated logarithm for U-statistics (Serfling (1980)),

$$n^{\frac{1}{2}}[p - H_n(\xi_p)] = O((\log \log n)^{\frac{1}{2}}) \text{ a.s.}.$$

Hence the first term on the right-hand side of (4.8) $\overset{a.s.}{\to}$ 0. The third term clearly $\to$ 0 also, and the second term $\overset{a.s.}{\to}$ to the right-hand side of (4.4). □

LEMMA 4.2. *For* N *corresponding to* (4.1), *with* $\hat{\zeta}_{pn}$ *satisfying* (4.3), *the r.v.'s* $\{Nd^2\}_{d>0}$ *are uniformly integrable.*

PROOF. By the proof of Lemma 3.2 of Bickel and Yahov (1968), it suffices to prove

$$\sum_{r=1}^{\infty} \sup_{0<d<d_0} P\{Nd^2 > r\} < \infty , \tag{4.9}$$

for some $d_0$. Let us write

$$P\{Nd^2 > r\} \leq P\{N > [r/d^2]\}$$

$$\leq P\{W_{[r/d^2],b([r/d^2])} - W_{[r/d^2],a([r/d^2])} \geq 2d\} . \tag{4.10}$$

By routine but tedious arguments (see Choudhury (1984), pp. 32-36, for details), one can bound the right-hand side of (4.10) by a function of r and d which is finitely summable in r uniformly in $d < d_0$, for sufficiently small $d_0$. □

LEMMA 4.3. *Let* $U_n$ *be the U-statistic based on kernel* $h(x_1,\ldots,x_n)$ *and a random sample* $X_1,\ldots,X_n$, $n \geq m$. *Let* $Eh^2 < \infty$ *and* $\eta_1 > 0$, *where*

$$\eta_1 = \mathrm{Var}\{E\{\textstyle\sum_A h(X_{i_1},\ldots,X_{i_m})/m!\,|X_1\}\} , \tag{4.11}$$

*and* $\sum_A$ *denotes summation over all permutations of* $(1,\ldots,m)$. *Let* $\{N_\Delta\}$ *be positive integer-valued r.v.'s and* $\{a_\Delta\}$ *positive constants, such that*

(4.12) $a_\Delta \to \infty$ *as* $\Delta \to \Delta_0$

*and*

(4.13) $N_\Delta / a_\Delta \to C_0$ *in probability, as* $\Delta \to \Delta_0$,

*with* $C_0$ *a finite constant. Then*

(4.14) $N_\Delta^{\frac{1}{2}}(U_{N_\Delta} - Eh) \xrightarrow{d} N(0, m^2\eta_1)$ *as* $\Delta \to \Delta_0$.

PROOF. Put

(4.15) $$U_n = \hat{U}_n + R_n,$$

where

$$\hat{U}_n = \sum_{i=1}^{n} E\{U_n | X_i\} - (n-1)Eh.$$

Then by Geertsema (1970) (or see Serfling (1980), p. 189), we have

(4.16) $$n^{\frac{1}{2}} R_n \xrightarrow{a.s.} 0, \quad n \to \infty.$$

Let $\varepsilon > 0$ be given. Then

$$P\{N_\Delta^{\frac{1}{2}} |R_{N_\Delta}| > \varepsilon\} \le P\{\sup_{k \ge C_0 a_\Delta / 2} k^{\frac{1}{2}} |R_k| > \varepsilon\} + P\{N_\Delta < C_0 a_\Delta / 2\}.$$

As $\Delta \to \Delta_0$, the first term on the right tends to 0 by virtue of (4.12)

and (4.16), and the second term tends to 0 by (4.13). Hence we obtain

$$N_\Delta^{\frac{1}{2}} R_{N_\Delta} \xrightarrow{p} 0, \text{ as } \Delta \to \Delta_0. \tag{4.17}$$

Also, by the classical CLT, we have

$$n^{\frac{1}{2}}(\hat{U}_n - Eh) \xrightarrow{d} N(0, m^2\eta_1) . \tag{4.18}$$

Thus, by the Doeblin-Anscombe CLT for randomly indexed sums (see Chow and Teicher (1978), p. 317), we have

$$N_\Delta^{\frac{1}{2}}(\hat{U}_{N_\Delta} - Eh) \xrightarrow{d} N(0, m^2\eta_1), \text{ as } \Delta \to \Delta_0. \tag{4.19}$$

Combining (4.17) and (4.19), the desired (4.14) follows. □

PROOF OF THEOREM 4.1. (a) By Lemma 4.1, i.e., by the convergence (4.4), we can easily deduce that $N(d)$ is finite a.s., that $N(d)$ is nondecreasing and $\to \infty$ a.s. as $d \to 0$, and finally (by monotone convergence) that $E\,N(d) \to \infty$ as $d \to 0$.

(b) Noting that

$$W_{N,b(N)} - W_{N,a(N)} \leq 2d < W_{N-1,b(N-1)} - W_{N-1,a(N-1)} ,$$

we have by the convergence $N \to \infty$ as $d \to 0$ and again the convergence (4.4) that

$$\lim_{d\to 0} N d^2 = [\Phi^{-1}(1-\alpha)]^2 m^2 \zeta_p / h_F^2(\xi_p) \text{ a.s.} \tag{4.20}$$

(c) It is readily seen that

$$\lim_{d\to 0} P\{W_{N,a(N)} \le \xi_p \le W_{N,b(N)}\} = \lim_{d\to 0} P\{N^{\frac{1}{2}}|H_N(\xi_p)-p| \le \Phi^{-1}(1-\alpha)m\hat{\zeta}_{pN}^{\frac{1}{2}}\},$$

and thus claim (c) follows via Lemma 4.3 applied to the U-statistic $H_n(\xi_p)$ and Slutsky's Theorem.

(d) Since by Lemma 4.2 the r.v.'s $\{Nd^2\}_{d>0}$ are uniformly integrable, the convergence (4.20) holds also in expectation. Finally, $E(N) < \infty$, $d > 0$, since the uniform integrability also implies $\sup_{d\to 0} E\{Nd^2\} < \infty$.

This completes the proof. □

REMARKS. (i) As noted previously, the parameter $\zeta_p$ is distribution-free only in exceptional cases, so that in typical applications a strongly consistent estimator is needed. Such an estimator is provided in Section 5.

(ii) The *asymptotic relative efficiency as* $d \to 0$ of a sequential fixed-width confidence interval T relative to another such procedure S may be taken as

$$e(T,S) = \lim_{d\to 0} E(N_S)/E(N_T).$$

For procedures satisfying Theorem 4.1, we obtain e(T,S) immediately from part (d) of the theorem. In particular, for the generalized Hodges-Lehmann location estimators $HL_{(m)}$ corresponding to $H_n^{-1}(\frac{1}{2})$ for the kernel $h(x_1,\ldots,x_m) = (x_1 + \ldots + x_m)/m$, we have the formula:

$$(4.21) \qquad e(HL_{(m)}, \bar{X}) = f^2_{\bar{X}_m}(\xi_{1/2})\sigma_F^2/m^2\zeta_{1/2}(F) ,$$

where $f_{\bar{X}}(\cdot)$ denotes the density function of the r.v. $\bar{X}_m = (X_1 + \ldots + X_m)/m$, $\sigma_F^2$ denotes the variance of the df F, $\zeta_{1/2}(F) = \mathrm{Var}_F\{F_0^{(m-1)}(X)\}$, and $F(x) = F_0(x - \xi_{1/2})$ with $F_0$ symmetric about 0. Values of (4.21) for several choices of F and m = 1,2,...,5 are as follows.

| | m | | | | |
|---|---|---|---|---|---|
| F | 1 | 2 | 3 | 4 | 5 |
| Normal | .637 | .955 | .981 | .989 | .993 |
| Uniform | .333 | 1.000 | .849 | .906 | .919 |
| Logistic | .822 | 1.097 | 1.103 | 1.083 | 1.077 |
| Laplace | 2.000 | 1.500 | 1.321 | 1.238 | 1.190 |

5. Estimation of the nuisance parameter $\zeta_p$. We note that $m^2\zeta_p$ is the asymptotic variance parameter of the U-statistic based on the kernel

$$(5.1) \qquad g(x_1,\ldots,x_m) = \mathbb{I}\{h(x_1,\ldots,x_m) \leq H_F^{-1}(p)\}.$$

Sen (1981), §3.7, for example, gives methodology for construction of strongly consistent estimators for the asymptotic variance parameters of U-statistics. However, these methods assume that the kernel of the U-statistic is completely known. In the present case, we have in (5.1) a kernel involving an unknown parameter $\xi_p = H_F^{-1}(p)$. Consequently, we develop an estimator by a different method.

Specifically, we introduce a family of kernels,

(5.2) $$K(x_1,\ldots,x_{2m-1},\Delta),\ (x_1,\ldots,x_{2m-1}) \in \mathbb{R}^{2m-1},$$

indexed by $\Delta$, such that

$$\zeta_p = EK(X_1,\ldots,X_{2m-1},\xi_p)\ .$$

We denote by $U_n(\Delta)$ the U-statistic based on the kernel in (5.2). Then a natural estimator of $\zeta_p$ is given by $U_n(\xi_p)$. Since, however, $\xi_p$ is unknown, we substitute its estimator $\hat{\xi}_{pn} = H_n^{-1}(p)$, arriving at the estimator

(5.3) $$\hat{\zeta}_{pn} = U_n(\hat{\xi}_{pn}).$$

THEOREM 5.1. *If* $H_F$ *is continuous at* $\xi_p$, *then a.s.* $\hat{\zeta}_{pn} \to \zeta_p$, $n \to \infty$.

PROOF. Define

$$J(x_1,\ldots,x_m,y) = (m!)^{-1} \sum \mathbb{I}\{h(x_{i_1},\ldots,x_{i_m}) \le y\},$$

where the sum is over permutations $(i_1,\ldots,i_m)$ of $(1,\ldots,m)$, and

$$G(x,y) = \int\ldots\int J(x_1,\ldots,x_{m-1},x,y) \prod_{i=1}^{m-1} dF(x_i).$$

Then the parameter $\zeta_p$ may be expressed as

(5.4) $$\zeta_p = \mathrm{Var}_F\{G(X,\xi_p)\} = E_F\{G^2(X,\xi_p)\} - p^2.$$

Defining

$$\theta(y) = \int G^2(x,y)dF(x) - p^2$$

and

$$K(x_1,\ldots,x_{2m-2},x,y) = J(x_1,\ldots,x_{m-1},x,y)J(x_m,\ldots,x_{2m-2},x,y) - p^2$$

we have

$$\theta(y) = \int\ldots\int K(x_1,\ldots,x_{2m-1},y) \prod_{i=1}^{2m-1} dF(x_i)$$

and

$$\zeta_p = \theta(\xi_p).$$

Noting that $K(x_1,\ldots,x_{2m-1},y)$ is monotone in the argument y, and that this function is continuous at $y = \xi_p$ with probability 1 with respect to the probability measure $\prod_{i=1}^{2m-1} dF(x_i)$, we obtain by the monotone convergence theorem that the function $\theta(y)$ is continuous at $y = \xi_p$.

Now let $\varepsilon > 0$ be given and choose $\delta > 0$ such that $|\theta(y) - \theta(\xi_p)| < \varepsilon$ for $|y - \xi_p| \leq \delta$. By the monotonicity of the kernel $K(x_1,\ldots,x_{2m-1},y)$ in the argument y, and by strong convergence of $\hat{\xi}_{pn}$ to $\xi_p$ (recall (2.4)), we have that a.s.

$$\hat{\zeta}_{pn} = U_n(\hat{\xi}_{pn}) \in [U_n(\xi_p - \delta), U_n(\xi_p + \delta)] \tag{5.5}$$

for all n sufficiently large. By the almost sure convergence of U-statistics, the interval in (5.5) is a.s. contained in the interval

$[\theta(\xi_p - \delta) - \varepsilon,\ \theta(\xi_p + \delta) + \varepsilon]$ for all n sufficientlv large.

But this latter interval is contained in $[\theta(\xi_p) - 2\varepsilon,\ \theta(\xi_p) + 2\varepsilon]$, i.e., in the interval $[\zeta_p - 2\varepsilon,\ \zeta_p + 2\varepsilon]$. □

REMARK. The estimator $\hat{\zeta}_{pn}$ given above requires $O(n^{2m-1})$ computational steps. In the case that O(n) computational ease is desired, one can use instead the estimator

$$\tilde{\zeta}_{pn} = \frac{1}{\left[\frac{n}{2m-1}\right]} \{K(X_1,\ldots,X_{2m-1},\hat{\xi}_{pn}) + K(X_{2m},\ldots,X_{4m-2},\hat{\xi}_{pn}) + \ldots\},$$

which is also strongly consistent but less efficient than $\hat{\zeta}_{pn}$. □

## REFERENCES


Bahadur, R.R. (1966), "A note on quantiles in large samples," *Ann. Math. Statist.*, 37, 577-580.

Bickel, P.J. and Lehmann, E.L. (1979), "Descriptive statistics for nonparametric models. IV. Spread," in *Contributions to Statistics. Hajek Memorial Volume*, 33-40 (ed. by J Jurečková), Academia, Prague.

Bickel, P.J. and Yahav, J.A. (1968), "Asymptotically optimal Bayes and minimax procedures in sequential estimation," *Ann. Math. Statist.* 39, 442-456.

Choudhury, J. (1984), *Sequential Fixed-Width Confidence Intervals Based on Generalized Order Statistics, and a Study of Generalized Hodges-Lehmann Location Estimators.* Unpublished Ph.D. dissertation. Department of Mathematical Sciences, Johns Hopkins University.

Chow, Y.S. and Teicher, H. (1978), *Probability Theory: Independence, Interchangeability, Martingales*, Springer-Verlag, New York.

Geertsema, J.C. (1970), "Sequential confidence intervals based on rank tests," *Ann. Math. Statist.*, 41, 1016-1026.

Ghosh, J.K. (1971), "A new proof of the Bahadur representation of quantiles and an application," *Ann. Math. Statist.*, 42, 1957-1961.

Hoeffding, W. (1963), "Probability inequalities for sums of bounded random variables," *J. Amer. Statist. Assoc.*, 58, 13-30.

Huber, P.J. (1964), "Robust estimation of a location parameter," *Ann. Math. Statist.*, 35, 73-101.

Janssen, P., Serfling, R. and Veraverbeke, N. (1984), "Asymptotic normality for a general class of statistical functions and applications to measures of spread," *Ann. Statist.*, 12, 1369-1379.

Maritz, J.S., Wu, M., and Staudte, Jr., R.G. (1977), "A location estimator based on a U-statistic," *Ann. Statist.*, 5, 779-786.

Sen, P.K. (1981), *Sequential Nonparametrics*, Wiley, New York.

Serfling, R.J. (1980), *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

Serfling, R.J. (1984), "Generalized L-, M- and R-statistics," *Ann. Statist.*, 12, 76-86.

Serfling, R.J. and Thornton, D.H. (1981), "An extension of Bahadur's representation of sample quantiles, with application to versions of the Hodges-Lehmann location estimator," Tech. Rept. No. 352 (ONR Technical Report No. 81-9), Department of Mathematical Sciences, Johns Hopkins University, Baltimore, Maryland.

Theil, H. (1950), "A rank-invariant method of linear and polynomial regression analysis, III," *Proc. Kon. Ned. Akad. v. Wetansch. A*, 53, 1397-1412.

J. Choudhury
Department of Computer Science, Mathematics, and Statistics
University of Baltimore
Baltimore, MD 21218

R. J. Serfling
Department of Mathematical Sciences
The Johns Hopkins University
Baltimore, MD 21218

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

| 1. REPORT NUMBER | 2. GOVT ACCESSION NO. | 3. RECIPIENT CATALOG NUMBER |
|---|---|---|
| ONR No. 85-4 | | |

| 4. TITLE | 5. TYPE OF REPORT & PERIOD COVERED |
|---|---|
| Generalized Order Statistics, Bahadur Representations, and Sequential Nonparametric Fixed-Width Confidence Intervals | Technical Report |
| | 6. PERFORMING ORGANIZATION REPORT NO. Technical Report No. 445 |

| 7. AUTHOR(s) | 8. CONTRACT OR GRANT NUMBER(s) |
|---|---|
| J. Choudhury and R.J. Serfling | ONR No. N00014-79-C-0801 |

| 9. PERFORMING ORGANIZATION NAME AND ADDRESS | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
|---|---|
| Department of Mathematical Sciences<br>The Johns Hopkins University<br>Baltimore, Maryland 21218 | |

| 11. CONTROLLING OFFICE NAME & ADDRESS | 12. REPORT DATE |
|---|---|
| Office of Naval Research<br>Statistics and Probability Program<br>Arlington, Virginia 22217 | October, 1985 |
| | 13. NUMBER OF PAGES 28 |

| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | 15. SECURITY CLASS (of this report) Unclassified |
|---|---|
| | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from report.

18. SUPPLEMENTARY NOTES

19. KEY WORDS

order statistics, Bahadur representation, sequential, nonparametric, fixed-width confidence intervals

20. ABSTRACT

Let $X_1,\ldots,X_n$ be an iid sample from df F, let $H_F$ be the df of $h(X_1,\ldots,X_m)$, based on a given "kernel" $h(x_1,\ldots,x_m)$, and consider confidence interval estimation of a parameter of the form $H_F^{-1}(p)$. This paper introduces confidence intervals formed by a pair of "generalized order statistics," develops Bahadur-type representation theory for those order statistics, and constructs corresponding sequential fixed-width confidence interval procedures. Previous work of Bahadur (1966) and Geertsema (1970) is sharpened and extended.

END

FILMED

2-86

DTIC